GROWTH EXTERNALITIES ON THE ENVIRONMENTAL QUALITY INDEX OF EAST JAVA INDONESIA, SPATIAL ECONOMETRICS MODEL OF STIRPAT[*]

Rahma Fitriani[1], Herman Cahyo Diartho[2], and Septya Hadiningrum[3]

[1]Department of Statistics, University of Brawijaya, Indonesia, rahmafitriani@ub.ac.id
[2]Department of Economics Development, University of Jember, Indonesia, hermancahyo.feb@unej.ac.id
[3]Department of Statistics, University of Brawijaya, Indonesia, septya03@gmail.com

Abstract

East Java has shown strong economic growth, which negatively affects its environmental quality. Analysis of the functional relationship between economic growth and environmental quality is important to direct the growth without further deteriorate the environmental quality in this area. It is assumed that growth produces some externalities on environmental quality. The spread of technological information, economic productivity, population growth or investment, can be the source of the growth externalities. The objective of this study is to test the significance of the involved growth externalities on East Java's environmental quality. Using spatial data, the externalities are accommodated in a spatial version of the STIRPAT model. It is estimated using per city/regency 2015 data. The analysis indicates that local density, local agricultural productivity, neighboring density, and neighboring mining activity significantly affect the local environmental quality. The latter two are the main sources of the growth externalities.

Keywords: environmental quality, externalities, spatial econometrics, STIRPAT.

1. Background and Objectives of the Study

Strong economic growth is a catalyst for regional development. Unfortunately, the growth invites some negative consequences for the quality of environment. Several studies on environmental economics have explored the functional relationship between economic growth and environmental quality. The form of the relationship serves as a basis to choose which factors can be controlled or directed, such that the growth can be maintained without further deteriorate the environmental quality.

Recent studies (Ertur & Koch, 2007; Fingleton & López-Bazo, 2006; Tian et al., 2010) indicate that the economic performance of one region is partly affected by the economic performance and social condition of its surrounding reigions. This interdependence implies that the local characteristic is not the only factor that determines regional economic activity. It depends also on the characteristics and the activities of the nearby regions. In that case, the activities affect the local as well as the neighbouring regions environmental quality. It is an indication of externalities created by economic growth. The increase in population size, productivity, and investment or the spread of technological information are equally possible to be the source of the growth externalities.

East Java, is one of Indonesia's provinces which performs a steady and strong economic growth. In the fourth quartal of 2017, its economic growth is above the national economic growth (BI, 2017). The economic structure has been dominated by agriculture (forestry and fishery), manufacturing and wholesale – retail trade (BPS, 2017). The province consists of 38 cities/regencies. The interaction between those cities/regencies depends on their geographical location and condition. Nearby cities/regencies, due to their similar geographical conditions, have similar economic activities (see Figure 1). The western part of the province is mountain area with mining potential, the fertile central and southern areas are dedicated to agriculture, whereas the majority of industrial activities are located in the low region in the north. The growth invites some consequences on environmental quality. In 2016, the index of East Java's environmental quality is 58.98 (100 for perfect quality), which is below the national quality index of 65.73 (KLH, 2016).

Figure 1: Map of dominant economic activity of cities/regencies in East Java (based on each city/regency's GDP by Sector 2015).

The relation between environmental quality and human activity is defined by IPAT equation (Ehrlich & Holdren, 1971). The equation defines that environmental quality is the product of Population size, Affluence, and Technology ($I = P \times A \times T$). Affluence can be represented by consumption or production. Since it is an identity ($T = I/(P \times A)$), it will be useless for measuring the magnitude of the effect of every factor on the environmental quality or testing its significance. For that purpose, York et al. (2003) develop a stochastic version of the equation, namely STIRPAT (Stochastic Impacts by Regression on Population, Affluence, and Technology). In this model, environmental quality is a function of Population size, Affluence (economic productivity) and Technology. Each model's parameters represent the magnitude of the effect of each factor on environmental quality. Within the context of econometrics, the model's parameters are estimated and tested for their significance.

Several studies analysed the effect of human activities (e.g. population size and structure, GDP, and consumption) on the environmental quality based on the STIRPAT model. Shahbaz et al. (2015), Zhu & Peng (2012) specifically estimated the model's parameters using time series data. The use of time series data mostly emphasizes the dynamic relationship between human activities and environmental quality. However, since the environmental quality naturally varies or shows a certain pattern across locations, extending the observation to more than one location will provide additional information about location variability or spatial pattern of the environmental quality. Time series of environmental quality across locations can be treated as a set of panel data (see Bargaoui et al. (2014) or Liddle (2013)), in which the causal relationship between the human activities and the environmental quality is analyzed within the setting of panel model. In the panel model, the estimated parameters will be adjusted for the possible difference characteristic among locations. But since the panel model does not explicitly use the spatial configuration of the observed locations, the causal relationship only holds locally. Wang et al. (2016) start to attach a spatial reference to the observed CO_2 emission of every province in China from 1995 to 2011. It is used only to calculate Moran's I statistics to test the significance of spatial autocorrelation of CO_2 emission among the China's 30 provinces. It is not utilized to capture the spatial dependence of CO_2 emission in the model. Thus even though the test confirm that the CO_2 emissions are spatially auto correlated across provinces, their developed model is not capable to capture the effect of one province's activity to the CO_2 emission of its nearby location.

The importance of spatial dimensions in environmental studies has been discussed by some authors (see Anselin (2001); Bockstael (1996)). The environmental quality in itself naturally has a spatial pattern (e.g. the spread of contaminated water or the diffusion of air pollution. Furthermore, it is common that regions interact with each other. They trade, exchange labor and technology, or expand capital. Ertur & Koch (2007) define this phenomenon as growth externalities. The nearby regions naturally will have similar economic activities leading to a more intense interaction, such that it is plausible to assume that they have suffered similar environmental problems. It would create a more apparent spatial pattern of the environmental quality, which is reflected by the significant spatial autocorrelation of the chosen indicator of environmental

quality. Regardless of the natural spatial pattern, Anselin (2001) points out other problems in most of the environment studies, which can be settled in the spatial modeling framework. It is the need to combine data from different sources based on different sample designs, such that the scale mismatch is unavoidable. The indicator of environment quality (i.e. carbon emission) is measured in several locations as point data, whereas the economic indicators are commonly presented as regional data. As a consequence the observations tend to show spatial dependencies or spatial heterogeneity, which needs to be accommodated in the modeling strategy.

More recent studies have implemented the spatial modeling framework for different purposes. Some focus on the hypothesis regarding the existence of the inverted U – shape of Environmental Kuznets Curve (EKC) using spatial econometrics model (Hao et al., 2018; W. Wang & Yu, 2015). They accommodate the spatial dependency in the model in order to have more meaningful estimated parameters of EKC, since applying a classic linear regression would only lead to the violation of independence errors. Some studies (Roberts, 2011; Videras, 2014) also include the spatial dimension to develop their STIRPAT model. The spatial version of STIRPAT in Roberts (2011) takes into account the spatial dependencies among the environmental impact as well as the spatial dependencies among the disturbance. It is estimated based on data at a lower administrative level as a unit of study (local) instead of data at the national level (global). The study focuses on showing that the relation between the environmental impact and its determining factors at the local level might be different from the relation at the global level. Whereas Videras (2014) uses a geographically weighted regression (GWR) to explore the issue of spatial heterogeneity in the STIRPAT framework. When the spatial data are available in a certain period of time (i.e. monthly, yearly), a spatial panel version of STIRPAT can be estimated. This approach is used by Liu et al. (2014), based on the spatial (provincial) panel data of carbon emissions (and it predictors) from 2010 – 2006. Within this setting, the spatial dependencies and heterogeneity are accommodated in the model, such that the effect of the predictors on the carbon emissions can be estimated more efficiently.

In general, the accommodation of the spatial dimension in those studies has been triggered by various research questions. However, they have not fully explored the nature of the involved growth externalities and how these externalities affect environmental quality. Geographically, the externalities' effect is captured by the spatial pattern (in terms of the spatial autocorrelation) of the environmental impact. Within the context of spatial modeling, the nature of externalities can be determined by analyzing which factors of the surrounding locations that significantly affect the environmental quality. It can be the neighbourhood environmental condition, neighbourhood growth, and productivity or the neighbourhood unobserved factors. The similar spatial pattern between the environmental quality and the clusters of economic activities indicates that the economic activities of the neighbouring locations play a certain role in shaping the local environmental quality.

For the case of East Java's environmental quality, Fitriani & Syukrilla (2017) have accommodated the externalities by estimating a spatial econometric model of STIRPAT. However, they failed to show the significance of economic productivity on

environmental quality. It is suspected that the use of overall GDP does not represent well enough the economic activity, in relation to the environmental quality. GDP by sectors should be used instead, by considering that different economics activity might affect the environment differently.

The objective of this study is to identify the involved growth externalities on East Java's environmental quality. It can be done by estimating the coefficients of the spatial version of STIRPAT, in which the local environmental quality index (EQI) is a function of neighbouring EQI, local as well as well as neighbouring regions' population density, investment, and GDP by sectors. It is followed by the hypothesis testing regarding the significance effect of each predictor on the EQI. The test is focused on the coefficients (of elasticity) of the neighbouring factors representing externalities. The estimated coefficients will be will be useful as references for policy makers to redirect the growth into a more environmentally friendly development in this area.

2. Research Data and Methodology

2.1   Research Data

The spatial version of STIRPAT is estimated using 2015 spatial data of 38 cities/regencies of East Java. The data consist of the observed values of the following variables:

- Environmental quality Index – EQI, of all 38 cities/regencies $I_i, i = 1,2,...,38$. A higher index indicates a better environment quality.
- Density (in 1000 persons/km^2) of all cities/regencies. Density is used instead of population size in order to accommodate the area of each city/regency. It is used as a proxy for $P_i, i = 1,2,...,n$.
- Per sector real GDP (in billion rupiahs) as the measure of productivity (A)
  - Real GDP per capita of the Agricultural Sector of each city/regency, $Ag_i, i = 1,...,38$
  - Real GDP per capita of the Mining Sector of each city/regency, $Min_i, i = 1,...,38$
  - Real GDP per capita of the Industrial Sector of each city/regency, $In_i, i = 1,...,38$
  - The fund invested for infrastructure (in trillion rupiahs) of all cities/regencies, $K_i, i = 1,...,38$.

2.2   Methodology

Defining the spatial version of STIRPAT
In order to analyze the relation between environment quality and its driving factors, IPAT identity has been modified into a stochastic model Stochastic Impacts by Regression on Population, Affluence, and Technology (STIRPAT) (York et al., 2003). The model assumes that environmental quality is a function of Population (P) and Affluence (A). Those two factors can be explicitly measured by population size and per capita productivity respectively. In the contrary, there is no agreement regarding the exact measure of Technology (T). Some (York et al., 2003) argue further that Technology (T) also depends positively on Population (P) and Affluence (A). Thanks

to the stochastics nature of the model, this problem can be solved by including T in the error terms.

York et al. (2003) define STIRPAT as:

$$I = aP^bA^cT^de \qquad (1)$$

in which each of the coefficients (b, c, d) represents the elasticity between the corresponding factor and the environmental quality, and e represents the random errors that independently and identically distributed. For the estimation purpose, (1) can be represented as an additive regression model:

$$\ln I_i = \ln a + b \ln P_i + c \ln A_i + d \ln T_i + e_i$$

$$= \beta_0 + b \ln P_i + c \ln A_i + d \ln T_i + e_i,$$

$$i = 1, 2, \cdots, n \qquad (2)$$

or

$$\ln I_i = \ln a + b \ln P_i + c \ln A_i + e_i$$

$$= \beta_0 + b \ln P_i + c \ln A_i + e_i,$$

$$i = 1, 2, \cdots, n \qquad (3)$$

This study modifies the model into a spatial version, in order to accommodate the growth externalities by providing spatial reference for each of the cross sectional data used in the estimation process. When a set of spatial data is used, it is possible to measure the growth externalities in terms of the local growth effect on the neighbouring locations or vice versa. The key to capture the externalities is in the definition of neighbouring locations. For that purpose, using $i = 1, \dots, n$ as a unit location index, the model uses a spatial lag operator. It defines a weighted average of the variable under study at 'neighbouring' locations. Here, 'the neighbours' of location i are locations that are significantly influential to location i or having intensive interaction with location i. They might be locations that share borders with location i or within a certain distance from location i. In this case, the neighbors correspond to the nearby locations.

Each location neighbours are defined by introducing a set of $n \times n$ spatial weight $(w_{ij}, i = 1, \cdots, n, j = 1, \cdots, n)$. It plays a role to differentiate the spatial model from the non-spatial one. If locations i and j are neighbours, then w_{ij} will be positive, zero otherwise. Since location i cannot be its own neighbour, then w_{ii} for $i = 1, \cdots, n$ are zeros. Elhorst (2014b) or Anselin (2013) presents some options to define the weight (contiguity based or distance based). This study adopts the contiguity concept, such that:

$$w_{ij} = \frac{c_{ij}}{c_{i.}}, c_{ij} = \begin{cases} 1, & \text{if location } i \text{ and } j \text{ are shared a border} \\ 0, & \text{otherwise} \end{cases}, c_{i.} = \sum_{j=1}^{n} c_{ij}. \qquad (4)$$

After defining a proper spatial weight matrix, the spatial version of STIRPAT is developed such that it is possible to measure how much the change of the

neighbourhood population size, the neighbourhood productivity, and the neighbourhood adopted technology affecting the local environmental quality, in addition to the effects which are created from the change in the local factors. This modeling concept falls within the spatial econometrics framework. There are some possible spatial econometrics models. Each of them is formed based on the assumed source of externalities, which affect the local response, namely: the neighbourhood response, the neighbourhood predictors and the neighbourhood unobserved factors. Detail regarding the classification of the spatial econometrics models has been discussed in Lesage & Pace (2009) and Elhorst (2014a).

The model which accommodates all of the possible sources of externalities is General Nesting Spatial (GNS). The corresponding GNS spatial version of STIRPAT based on model (2) is:

$$\ln I_i = \beta_0 + \phi \sum_{j=1}^{n} w_{ij} \ln I_i + b \ln P_i + c \ln A_i + d \ln T_i + \rho \sum_{j=1}^{n} w_{ij} \ln P_j + \alpha \sum_{j=1}^{n} w_{ij} \ln A_j$$

$$+ \kappa \sum_{j=1}^{n} w_{ij} \ln T_j + e_i$$

$$e_i = \gamma \sum_{j=1}^{n} w_{ij} e_j + v_i,$$

$$i = 1,2,\cdots,n \tag{5}$$

or the following one based on model in (3) when T is included in the error terms:

$$\ln I_i = \beta_0 + \phi \sum_{j=1}^{n} w_{ij} \ln I_i + b \ln P_i + c \ln A_i + d \ln T_i + \rho \sum_{j=1}^{n} w_{ij} \ln P_j + \alpha \sum_{j=1}^{n} w_{ij} \ln A_j$$

$$+ e_i$$

$$e_i = \gamma \sum_{j=1}^{n} w_{ij} e_j + v_i,$$

$$i = 1, 2, \cdots, n \tag{6}$$

$\phi, \rho, \alpha, \kappa$ and γ in (5) or $\phi, \rho, \alpha,$, and γ in (6) are the parameters which measure the effect of externalities on the local environmental quality. Specifically, ϕ measures the effect of neighbourhood environmental quality, $\rho, \alpha,$ and $\kappa,$ measure the effect of the neighbourhood predictors (e.g. population, productivity, and technology), and γ measures the effect of the neighbourhood unobserved factors. In model (5) and (6), there is a chance that the error terms $e_i, i = 1, \ldots, n$ are spatially auto-correlated. Therefore, additional error terms that independently, identically and normally distributed are defined, namely $v_i, i = 1, \ldots, n$. Since each variable in (5) and (6) is

defined in its *In* form, its parameter is interpreted as a coefficient of elasticity of the local environmental quality due to the change of the corresponding variable.

The relation is still applicable for the level variables (without *In* transformation) such that the following models:

$$I_i = \beta_0 + \phi \sum_{j=1}^{n} w_{ij} I_i + bP_i + cA_i + dT_i + \rho \sum_{j=1}^{n} w_{ij}P_j + \alpha \sum_{j=1}^{n} w_{ij}A_j + \kappa \sum_{j=1}^{n} w_{ij}T_j + e_i$$

$$e_i = \gamma \sum_{j=1}^{n} w_{ij}e_j + v_i, \quad i = 1,2,\cdots,n \tag{7}$$

and

$$I_i = \beta_0 + \phi \sum_{j=1}^{n} w_{ij} I_i + bP_i + cA_i + dT_i + \rho \sum_{j=1}^{n} w_{ij}P_j + \alpha \sum_{j=1}^{n} w_{ij}A_j + e_i$$

$$e_i = \gamma \sum_{j=1}^{n} w_{ij}e_j + v_i, \quad i = 1,2,\cdots,n \tag{8}$$

hold as an alternative respectively for (5) and (6). They are the GNS spatial version of STIRPAT using level data. Each parameter in (7) and (8) is no longer interpreted as the coefficient of elasticity. It represents the marginal effect of environmental quality, due to the change of the corresponding predictor. However, the coefficient of elasticity can still be calculated based on the following relation:

$$\boldsymbol{E} = \boldsymbol{M} \times \frac{Average\ of\ Predictor}{Average\ Response} \tag{9}$$

M is the estimated marginal effect and *E* is the estimated coefficient of elasticity (Gujarati, 2003; Mankiw, 2014). When one or some of those parameters are not significant, GNS can be reduced to a simpler model that is nested in it (see Elhorst (2014a) for detail classification).

Defining the Spatial Version of STIRPAT in Terms of the Observed Variables
For the estimation of the spatial version of STIRPAT, the available research data have some limitation. Firstly, there is no proxy for measuring the level of adopted technology for each city/regency such that it is assumed that T is included in the error terms. Secondly, some cities/regencies have no real GDP of Mining Sector ($Min_i = 0$), such that the log transformation of this variable will be undefined. Therefore a GNS for the level variable by assuming that T is included in the error terms (model in (8)) is chosen as a starting point. The parameter which captures the spatial autocorrelation among the error terms represents the magnitude of externalities produced by the unobserved factors (e.g. adopted technology). Furthermore even though the estimated parameters are the marginal effect, they can be modified into the coefficients of elasticity, by applying relation in (9).

The model in terms of the observed variable can be defined as follows:

$$I_i = \phi \sum_{j=1}^{n} w_{ij}I_j + \beta_0 + bP_i + c_1 Ag_i + c_2 Min_i + c_3 In_i + dK + \rho \sum_{j=1}^{n} w_{ij}P_j + \alpha_1 \sum_{j=1}^{n} w_{ij}Ag_j$$

$$+ \alpha_2 \sum_{j=1}^{n} w_{ij}Min_j + \alpha_3 \sum_{j=1}^{n} w_{ij}In_j + \kappa \sum_{j=1}^{n} w_{ij}K_j + e_i$$

$$e_i = \gamma \sum_{j=1}^{n} w_{ij}e_j + v_i. \tag{10}$$

In (10) ϕ measures the effect of neighbourhood environmental quality, $\rho, \alpha_1, \alpha_2, \alpha_3$ and $\kappa,$ measure the effect of the neighbourhood predictors (e.g. density, agricultural sector productivity, mining sector productivity, industrial sector productivity and fund invested for infrastructure respectively), and γ measures the effect of the neighbourhood unobserved factors. The focus will be on the significance of those parameters. They can be used to reduce GNS into a simpler model and to answer the question regarding the involved externalities on East Java's environmental quality.

Estimating the Models' Parameters

The parameter estimation method depends on the problem emerged due to the setting of each model. Arbia (2014) provides a detailed discussion about the estimation method for each model. Generally, the setting of the model creates one or the combination of the following problems: heteroscedasticity, (spatial) autocorrelation among the error terms and endogeneity. The number of problems that the model has depends on its complexity. Due to their most complex setting, GNS and SAC (Spatial Autoregressive Confused) have all of those three problems. In this situation one of the following methods is applicable: Maximum Likelihood (ML), Generalized Spatial Two Stage Least Squares (GS2SLS), or Lee's Instrumental Variable (LIV). SDEM (Spatial Durbin Error Model) and SEM (Spatial Error Model) suffer from heteroscedasticity and (spatial) autocorrelation among the error terms, such that ML or Feasible GLS (FGLS) is considered appropriate. ML or Two – Stage Least Squares (2SLS) can be applied for SAR (Spatial Autoregressive) and SDM (Spatial Durbin Model) because they only have endogeneity problem. Because SLX (Spatial Lag of X) is free from those problems, OLS is applicable to estimate its parameters.

Choosing the Best Model

The best model is the model which has already accommodated all possible sources of externalities. It can be GNS or other simple model that nested in GNS. It starts with GNS in (10). It is considered as a candidate of the best model if all the sources of externalities create significant effect on EQI. Otherwise one source of externalities with insignificant effect is eliminated, leading to one of the following models: SAC, SDEM or SDM. The elimination of one more source of externalities is repeated if the relevant variables do not show significant effect on EQI. It leads to one of the following models which only have one source of externalities, namely: SAC, SDEM or SDM.

The model that indicates the significant effect(s) of the remaining neighbourhood variable(s) is considered as the model that well capturing the externalities on the

environmental quality. On the other hand, if up to the last model there is no evidence regarding the significance of the remaining neighbourhood variable, it is an indication that the EQI is only affected by the local factors.

Within the context of spatial econometrics modeling, without explicitly accommodates the spatial externalities, if the effects are indeed significant, they will be captured by the model's error terms. In this case, the null hypothesis of no spatial autocorrelation among the model's residuals will be rejected. On the other hand the null hypothesis about spatial independence of residuals is accepted if the model has fully taken into account the spatial externalities. Therefore, to guarantee that the chosen model has well accommodated the spatial externalities, LM tests for no spatial autocorrelation among the model's residuals against specific (alternative) models are conducted.

Technical details regarding LM test can be found in Anselin (2013) or Arbia (2014). Instead of conducting a Moran test, both writers prefer this type of test. They argue that unlike the Moran test, which does not state how to model the spatial autocorrelation if the null hypothesis is rejected, the LM test explicitly defined SEM or SAR as the alternative model to take into account the spatial autocorrelation.

Estimating the Direct and the Indirect Effects

The definitions about the direct and indirect effects have been discussed thoroughly in Lesage & Pace (2009) and Arbia (2014). The magnitudes of those effects depend on the parameters of the model. They are calculated as the partial derivative of the expected local EQI with respect to each local or neighboring predictor, namely: density (P), agricultural productivity (Ag), mining productivity (Min), industrial productivity (In) and invested fund for infrastructure (K). The direct effect is derived as the effect of local activity (in terms of those predictors) on the local EQI, whereas the indirect effect is the effect of neighboring activity on the local EQI.
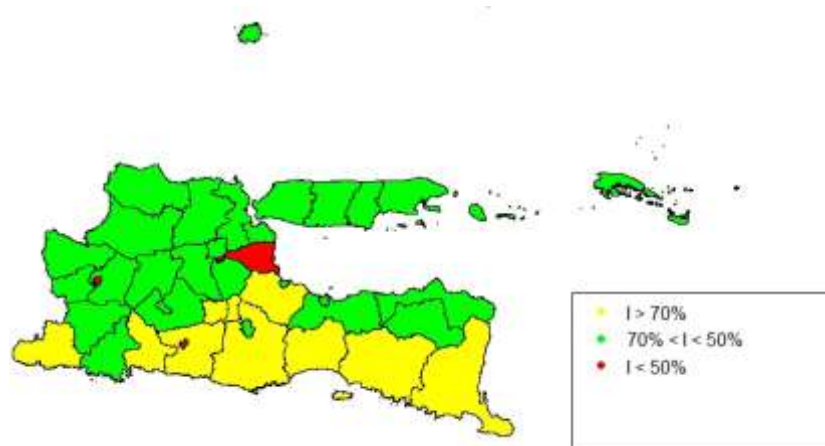

Figure 2: The map of Environmental Quality Index of Each City/Region of East Java Indonesia, 2015 Data.

3. Results and Discussion

Before discussing the estimation process and the significance of each factor on the environmental quality, it is important to understand the spatial pattern of the environmental quality index in this study area. The map of observed environmental quality index of each city/regency, from 2015 data is presented in Figure 2. The comparison between Figure 2 and the map of dominant economic activity in Figure 1 reveals that the cities/regencies with low environmental quality index are dominated by industrial activity. On the contrary, regions having high EQI are mainly dominated by agricultural activity. In Figure 2, the clustering of nearby cities/regencies with the same EQI categories indicates a positive spatial autocorrelation among the EQI of cities/regencies in East Java. This indication is confirmed by the result of Moran I which tests the spatial autocorrelation of EQI among East Java's cities/regencies. The test yields p value = 0.069, which is adequate support (at $\alpha = 10\%$) for the significance of the spatial autocorrelation.

The estimated parameters for all models are presented in Table 1. GNS is the starting point of the modeling. If one or more parameters which represent the effect of externalities are not significant, a simpler model can be chosen. The result in Table 1 indicates that the estimated parameters of neighbourhood EQI (ϕ), the neighbourhood unobserved factors (γ) and some of the neighbourhood s predictors ($\rho, \alpha_1, \alpha_2, \alpha_3, \kappa$) of GNS are not significant. Therefore SAC, SDEM or SDM can be fitted, by eliminating one type of the neighbourhood factor, as follows:
- **SAC**: It is fitted by eliminating the parameters of the neighbourhood predictors ($\alpha_1 = \alpha_2 = \alpha_3 = \kappa = 0$). The neighbourhood EQI and the neighbourhood unobserved factors are still in the model. However, the result in Table 1 shows that the estimated parameters for those variables (ϕ and γ) are not significant, such that one of those variables will be excluded in the next modeling stage.
- **SDEM**: It is fitted by eliminating the parameter of the neighbourhood EQI ($\phi = 0$). The result in Table 1 indicates that the estimated parameter of some of the neighbourhood predictors (ρ, α_2, κ) have significant effects on the local EQI. Only the unobserved factors which do not significantly affect the EQI, therefore in the next stage, these factors will be excluded.
- **SDM**: It is fitted by eliminating the parameter of the neighbourhood unobserved factors ($0 = \gamma$). The estimated parameters and the result of hypotheses testing in Table 1 show that all the remaining neighbourhood variables do not significantly affect the environmental quality. One of those neighbourhood variables will be excluded in the next stage.

The next stage of the modeling process is estimating the models which only include one type of neighbourhood variable, as follows:
- **SAR:** It is fitted by eliminating the parameters of the neighbourhood predictors ($\alpha_1 = \alpha_2 = \alpha_3 = \kappa = 0$) and the neighbourhood unobserved factors ($\gamma = 0$). However, after hypothesis testing (see Table 1), the remaining neighbourhood variable, namely the neighbourhood environmental quality does not show

significant effect (ϕ) on the local environmental quality.

- **SEM:** It is fitted by eliminating the parameters of the neighbourhood predictors ($\alpha_1 = \alpha_2 = \alpha_3 = \kappa = 0$) and the neighbourhood environmental quality ($\phi = 0$). The estimated parameter (γ) for the remaining neighbourhood variable, which represents the neighbourhood unobserved factors, is not significant.
- **SLX:** It is fitted by eliminating the parameter of neighbourhood environmental quality ($\phi = 0$) and the neighbourhood unobserved factors ($\gamma = 0$). The neighbourhood predictors are still in the model and the hypothesis testing indicates that some of the parameters ($\alpha_1, \alpha_2, \alpha_3, \kappa$) for these variables are significant.

Table 1: The Estimated Parameters of all models.

Parameter	Variable	GNS	SAC	SDEM	SDM	SAR/SLM	SEM	SLX
β_0	Intercept	66.782 (**)	37.514 (**)	71.509 (**)	88.92	53.076 (**)	61.963 (**)	69.130 (**)
ϕ	Neighbourhood *I*	0.076	0.381		-0.302	0.134		
b	*P*	-0.002 (**)	-0.002 (**)	-0.002 (**)	-0.002	-0.001 (**)	-0.001 (**)	-0.002 (**)
c_1	*Ag*	0.001 (**)	0.001 (**)	0.001 (**)	0.001	0.001 (**)	0.001 (**)	0.001 (**)
c_2	*Min*	-1.39e-04	-2.15e-04	-1.61e-04	-2.37e-04	-1.37e-04	-2.02e-04	-1.744e-04
c_3	*In*	-1.26e-06	-4.93e-05	3.62e-06	1.82e-05	-2.25e-05	-6.19e-05	6.52e-05
D	*K*	0.297	0.162	0.295	0.286	0.035	0.068	0.205
ρ	Neighbourhood *P*	-0.007 (**)		-0.008 (**)	-0.008			-0.007 (*)
α_1	Neighbourhood *Ag*	-2.54e-04		-8.12e-05	7.27e-04			2.93e-04
α_2	Neighbourhood *Min*	-1.01e-03		-1.04e-03 (**)	-1.09e-03			-8.39e-04 (*)
α_3	Neighbourhood *In*	1.25e-04		1.05e-04	-7.80e-06			1.01e-06
κ	Neighbourhood *K*	0.111		0.172	0.471			0.326
γ	Neighbourhood unobserved factor	-0.478	-0.465	-0.408			-0.079	

- (**) significant at 5% of α, (*) significant at 10% of α.

Up to the second stage of modeling, the only model which shows the significance of the remaining neighbourhood variable(s) is SLX. The LM test is conducted to check the nature of the model's residual terms. No spatial autocorrelation among the residual terms indicates that all possible neighbourhood factors which affect the local

environmental quality (e.g. adopted technology) have been well accommodated in the model. It is used as the null hypothesis, which is tested against two alternative models, SEM and SAR. The results of the test for both models are presented in Table 2. The test does not reject the null hypothesis for both alternatives. It guarantees that SLX is the best model. Therefore, all research questions which are related to East Java's environmental condition are answered based on the SLX.

Table 2: The result of LM test for the spatial independency of the SLX's residual/terms.

	LM Statistics	P value
SEM for Alternative	1.286	0.2567
SAR for Alternative	0.972	0.3241

The estimated parameters for SLX are depicted in the last column of Table 1. The result of the hypothesis testing for each of SLX's parameters indicates that the local EQI is significantly affected by the local and neighbourhood density, the neighbourhood mining productivity, and the local agricultural productivity. The positive effects are produced by the local agricultural productivity, whereas the local density, neighbourhood density, and neighbourhood mining productivity create negative effect on the local EQI.

In the first part of this paper, it is formulated that the aim of this study is to test the significance of the growth externalities on the EQI of East Java. To achieve that objective, this discussion is focused on the significance of parameters that are related to the neighbourhood factors. Since SLX is considered as the best model, automatically the neighbourhood EQI, and the neighbourhood unobserved factors have already excluded from the model. It implies that they are not the source of externalities. Among the estimated parameters of SLX for the East Java EQI presented in the last column of Table 1, ρ and α_2 are the significant parameters. They measure the effect of neighbourhood density and neighbourhood mining activity respectively. Therefore, for the case of East Java's EQI the density (population) and mining activity are the main source of externalities.

Before discussing the policy implications of the result, it is necessary to analyze the interpretation of each coefficient which measures the effect of the change of one predictor on the EQI. In an ordinary regression model, the parameter of a certain predictor is the marginal rate of the response (for the level variables) or the coefficient of elasticity of the response (for the log variables), with respect to the change of that particular predictor, by holding other predictors constant. The change occurs at location i and it is assumed that the effect only takes place in location i, without affecting the situation in other locations. The complexity may arise when a spatial econometric model is used. When the change of a particular predictor takes place in location i, in addition the effect at the same location, the change will also affect other locations (the neighbourhood). Technically those effects are defined as direct effect and indirect effect, respectively.

The magnitudes of those effects depend on the parameters of the model. In order to understand how the change of each predictor in every location affecting the response (EQI), the GNS model for East Java's EQI in (10) is redefined specifically for SLX, such that:

$$I_i = \beta_0 + bP_i + c_1 Ag_i + c_2 Min_i + c_3 In_i + dK + \rho \sum_{j=1}^{n} w_{ij} P_j + \alpha_1 \sum_{j=1}^{n} w_{ij} Ag_j + \alpha_2 \sum_{j=1}^{n} w_{ij} Min_j + \alpha_3 \sum_{j=1}^{n} w_{ij} In_j + \kappa \sum_{j=1}^{n} w_{ij} K_j + e_i \qquad (11)$$

with its respective expected value:

$$E(I_i) = \beta_0 + bP_i + c_1 Ag_i + c_2 Min_i + c_3 In_i + dK + \rho \sum_{j=1}^{n} w_{ij} P_j + \alpha_1 \sum_{j=1}^{n} w_{ij} Ag_j$$

$$+ \alpha_2 \sum_{j=1}^{n} w_{ij} Min_j + \alpha_3 \sum_{j=1}^{n} w_{ij} In_j + \kappa \sum_{j=1}^{n} w_{ij} K_j$$

$$(12)$$

The following $n \times n$ matrices: $\boldsymbol{S}_P, \boldsymbol{S}_{Ag}, \boldsymbol{S}_{Min}, \boldsymbol{S}_{In}$ and $\boldsymbol{S}_K$ are the matrices of the partial derivative of the expected value of EQI with respect to each predictor, density (P), agricultural productivity (Ag), mining productivity (Min), industrial productivity (In) and invested fund for infrastructure (K) respectively. The ij^{th} element of each matrix is respectively defined as:

$$S_{P,ij} = \frac{\partial E(I_i)}{\partial P_j}, S_{Ag,ij} = \frac{\partial E(I_i)}{\partial Ag_j}, S_{Min,ij} = \frac{\partial E(I_i)}{\partial Min_j}, S_{In,ij} = \frac{\partial E(I_i)}{\partial In_j}, S_{K,ij} = \frac{\partial E(I_i)}{\partial K_j}$$

$$i = 1, \dots, n, j = 1, \dots, n$$

Each element measures the expected effect of the change of a particular predictor in location j on the EQI of location i.

According to the source of the change, the effects can be categorized as follows:
- **The Direct effect**: it is the effect on the response observed at location i when the change of a predictor takes place also at location i (S_{ii}). For East Java's EQI, based on SLX in (11), the magnitude of this effect is represented by the parameter of a local predictor, which are b, c_1, c_2, c_3 and d respectively for the change of density (P), agricultural productivity (Ag), mining productivity (Min), industrial productivity (In) and invested fund for infrastructure (K). The magnitude of this effect is constant across locations, $i = 1, \dots, n$, such that, it also represents the magnitude of **the Average Direct Effect (ADE)**.
- **The average total effect from location i**: it is the average effect of the response at location i due to the change of a particular predictor across location $j = 1, \dots, n$

$$n^{-1} \sum_{j=1}^{n} S_{ij}$$

Summing the effects from all locations $i = 1, \dots, n$, will define **The Average Total Effect (ATE).** For East Java's EQI, based on SLX in (11), the magnitude of this

effect is $(b + \rho), (c_1 + \alpha_1), (c_2 + \alpha_2), (c_3 + \alpha_3)$ and $(d + \kappa)$ respectively for the change of density (P), agricultural productivity (Ag), mining productivity (Min), industrial productivity (In) and invested fund for infrastructure (K).

- **The Average Indirect Effect (AIE)**: it is the total effect of the response across locations $i, i = 1, ..., n$, when the change of a predictor occurs elsewhere (location $j \neq i$). It can be calculated as:

$$AIE = ATE - ADE$$

For East Java's EQI, based on SLX in (11), the magnitude of this effect is $\rho, \alpha_1, \alpha_2, \alpha_3$ and κ respectively for the change of density (P), agricultural productivity (Ag), mining productivity (Min), industrial productivity (In) and invested fund for infrastructure (K).

All of those effects take place by assuming that only one predictor has changed its state, by holding the others constant. The definitions about the direct and indirect effects have been discussed thoroughly in Lesage & Pace (2009) and Arbia (2014). Since the model is estimated using level data, the estimated parameters of SLX are interpreted as the marginal rate of change. The equation (9) can be applied in order to change them into the coefficients of elasticity. The magnitude of each effect, in terms of the Marginal Effect and the Coefficient of Elasticity, is presented in Table 3.

Table 3: The Estimated Total Effects, Direct Effects and Indirect Effects, in terms of Marginal Effect and Coefficient of Elasticity.

		Marginal Effect				Coefficient of Elasticity		
		Total Effect	Direct Effect	Indirect Effect	Average	Total Effect	Direct Effect	Indirect Effect
Source of change	P	-0.0091	-0.0022	-0.0069	1829.6579	-0.2572	-0.0616	-0.1957
	Ag	0.0016	0.0013	0.0003	4288.4658	0.1045	0.0851	0.0194
	Min	-0.0010	-0.0002	-0.0008	1744.8896	-0.0273	-0.0047	-0.0226
	In	0.0001	0.0001	0.0000	10316.6368	0.0106	0.0104	0.0002
	K	0.5304	0.2048	0.3256	2.4257	0.0199	0.0077	0.0122
Response	***EQI***				65			

The following discussion regarding the magnitude of each significance effect in terms of coefficient elasticity is based on the estimated effects presented in Table 2. It is assumed that the change of EQI is triggered by the change of one of the predictors, which significantly affect the EQI, namely: density (P), agricultural productivity (Ag), and mining productivity (Min).

The increase in density (P) creates a significant effect on the local East Java's EQI (direct effect) as well as the neighbourhood EQI (indirect effect). When other factors are held constant, 1% increases in density decreases the local EQI by 0.062% on average and the EQI of other cities/regencies by 0.19% on average. This result indicates that in East Java indeed the more populated an administrative area is, the

environment faces more pressure. Locally, more land needs to be cleared for residential, more carbon is produced and more pollutant is disposed into the open water due to the more economic activities from its citizen. The mobility of labor, which resides in one city/regency but works in one of nearby regencies/cities, eventually creates similar problems in his working regency/city. However the percentage change of EQI is not as big as the percentage increase of density, directly and indirectly.

The increases in agricultural productivity (Ag) only create a significant local (direct) effect on East Java's EQI. In this case, when other factors are held constant, 1% increases in agricultural productivity (Ag) increases the EQI on average by 0.08%. The more agricultural activity leads to more land is preserved for green area. But since land is limited inside a certain administrative boundary, the activity only affects the local EQI, it does not produce any externalities on the EQI of the other cities/regencies. In East Java's case the local EQI is inelastic for the change of agricultural productivity.

On the other hand, even though it is not big in magnitude, the analysis suggests that mining productivity (Min) creates a certain impact on the EQI of surrounding cities/regencies of East Java. The significance of its indirect effect indicates that 1% increase of mining productivity leads to an average of 0.026% decrease on the EQI of other cities/regencies. The result indicates that even though the activity does not create significant effect locally, the pollutant produced by this activity has been carried by water or air, which is more impactful when it reaches certain radius beyond the local administrative boundary.

Regarding the objective of this study, density (P) and mining productivity (Min) are identified as the involved growth externalities on East Java's environmental quality. The effect of the change of those two factors might reach the nearby cities/regencies, and affect the EQI in those cities/regencies.

4. Conclusions

A spatial version of STIRPAT is developed in order to identify the dominant growth externalities on the environmental quality index (EQI) among 38 cities/regencies of East Java. The model is estimated based on a set of spatial data of each city/regency's EQI, density, density (P), mining productivity (Min), agricultural productivity (Ag), industrial productivity (In) and invested fund for infrastructure (K). A GNS model is the initial estimated model, which then proceed to a simpler model due to the insignificance of one or some parameters. SLX is chosen as the model which appropriately defines the EQI as a function of local as well as neighbourhood predictors. This study also derives the direct and the indirect effects for the change of each predictor on the EQI, such that the magnitude of the local, as well as the neighbourhood effects, can be measured. The SLX identifies that for East Java's case, density (P) and the mining productivity (Min) are the involved growth externalities that affect the EQI.

The model indicates that density (P) creates some negative impacts on the local as well as neighbourhood EQI, agricultural productivity (Ag) produces a positive effect

on the local EQI, and mining productivity (Min) creates a negative effect on the EQI of the surrounding cities/regencies. The highlight of this study is that in every city/regency the population density has reached the maximum environment carrying capacity, such that a slight increase in the population density leads to a worse local as well as neighbourhood environmental quality. But, the magnitude of the effects is less than the percentage of the increase in density. In terms of productivity by sector, more economic activity does not necessarily decrease the environmental quality. The result indicates that maintaining or preserving agricultural productivity can be a potential solution to have a better environment quality, even though the causal effect relationship only holds locally. Special attention must be given to cities/regencies with mining potential. There should be coordination among the particular cities/regencies with the surrounding cities/regencies in terms of economic tools (i.e. tax) to internalize the negative externalities of the mining activity.

References

Anselin, L. (2001). Spatial effects in econometric practice in environmental and resource economics. *American Journal of Agricultural Economics, 83*: 705–710.

Anselin, L. (2013). *Spatial econometrics: methods and models*. Springer Science & Business Media.

Arbia, G. (2014). *A primer for spatial econometrics: with applications in R*. Springer.

Bargaoui, S. A., Liouane, N., & Nouri, F. Z. (2014). Environmental impact determinants: An empirical analysis based on the STIRPAT model. *Procedia-Social and Behavioral Sciences, 109*(2): 449–458.

[BI] Bank Indonesia. (2017). *Laporan Perekonomian Indonesia tahun 2017*. Jakarta (ID): Bank Indonesia.

Bockstael, N. E. (1996). Modeling Economics and Ecology: The Importance of a Spatial Perspective. *American Journal of Agricultural Economics, 78*: 1168–1180.

[BPS] Badan Pusat Statistik. (2017). *Produk Domestik Regional Bruto Provinsi Jawa Timur Menurut Lapangan Usaha 2012-2016*. Jakarta (ID): Badan Pusat Statistik.

Ehrlich, P. R., & Holdren, J. P. (1971). Impact of population growth. *Science, New Series, 171*: 1212–1217.

Elhorst, J. P. (2014a). *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels, Dordrecht*. Springer Briefs in Regional Science.

Elhorst, J. P. (2014b). Spatial panel data models. In *Spatial econometrics* (pp. 37–93). Springer.

Ertur, C., & Koch, W. (2007). Growth, technological interdependence and spatial externalities: theory and evidence. *Journal of Applied Econometrics, 22*: 1033–1062.

Fingleton, B., & López-Bazo, E. (2006). Empirical growth models with spatial effects*. *Papers in Regional Science, 85*: 177–198.

Fitriani, R., & Syukrilla, W. (2017). Growth Externalities on the Environmental Quality Index of East Java Indonesia, Spatial Econometrics Model. *Proceeding Regional Statistics Conference*, 20–24. Denpasar (ID): International Statistical Institute.

Gujarati, D. N. (2003). *Basic econometrics*. Boston (US): McGraw Hill.

Hao, Y., Wu, Y., Wang, L., & Huang, J. (2018). Re-examine environmental Kuznets curve in China: Spatial estimations using environmental quality index. *Sustainable Cities and Society*, *42*: 498–511.

[KLH] Kementerian Lingkungan Hidup dan Kehutanan. (2016). *Indeks Kualitas Lingkungan Hidup Indonesia 2016*. Jakarta (ID): Kementerian Lingkungan Hidup dan Kehutanan.

Lesage, J. P., & Pace, R. K. (2009). *Introduction to spatial econometrics*. Boca Raton (US): CRC Press.

Liddle, B. (2013). Population, affluence, and environmental impact across development: Evidence from panel cointegration modeling. *Environmental Modelling & Software*, *40*: 255–266.

Liu, Y., Xiao, H., Zikhali, P., & Lv, Y. (2014). Carbon emissions in China: a spatial econometric analysis at the regional level. *Sustainability*, *6*: 6005–6023.

Mankiw, N. G. (2014). *Essentials of economics*. Cengage learning.

Roberts, T. D. (2011). Applying the STIRPAT model in a post-Fordist landscape: Can a traditional econometric model work at the local level? *Applied Geography*, *31*: 731–739.

Shahbaz, M., Loganathan, N., Sbia, R., & Afza, T. (2015). The effect of urbanization, affluence and trade openness on energy consumption: A time series analysis in Malaysia. *Renewable and Sustainable Energy Reviews*, *47*: 683–693.

Tian, L., Wang, H. H., & Chen, Y. (2010). Spatial externalities in China regional economic growth. *China Economic Review*, *21*.

Videras, J. (2014). Exploring spatial patterns of carbon emissions in the USA: a geographically weighted regression approach. *Population and Environment*, *36*: 137–154.

Wang, S., Fang, C., & Wang, Y. (2016). Spatiotemporal variations of energy-related CO2 emissions in China and its influencing factors: An empirical analysis based on provincial panel data. *Renewable and Sustainable Energy Reviews*, *55*: 505–515.

Wang, W., & Yu, J. (2015). Estimation of spatial panel data models with time varying spatial weights matrices. *Economics Letters*, *128*: 95–99.

York, R., Rosa, E. A., & Dietz, T. (2003). STIRPAT, IPAT and ImPACT: analytic tools for unpacking the driving forces of environmental impacts. *Ecological Economics*, *46*: 351–365.

Zhu, Q., & Peng, X. (2012). The impacts of population change on carbon emissions in China during 1978-2008. *Environmental Impact Assessment Review*, *36*: 1–8.